AFRL-AFOSR-JP-TR-2018-0027

Social Dynamics Modeling and Inference

Kwang-Cheng Chen
NATIONAL TAIWAN UNIVERSITY

03/29/2018
Final Report

Air Force Research Laboratory
AF Office Of Scientific Research (AFOSR)/ IOA
Arlington, Virginia 22203
Air Force Materiel Command

REPORT DOCUMENTATION PAGE

Form Approved
OMB No. 0704-0188

1. REPORT DATE *(DD-MM-YYYY)*	2. REPORT TYPE	3. DATES COVERED *(From - To)*
30-03-2018	Final	14 May 2014 to 13 May 2017

4. TITLE AND SUBTITLE
Social Dynamics Modeling and Inference

5a. CONTRACT NUMBER

5b. GRANT NUMBER
FA2386-14-1-4053

5c. PROGRAM ELEMENT NUMBER
61102F

6. AUTHOR(S)
Kwang-Cheng Chen

5d. PROJECT NUMBER

5e. TASK NUMBER

5f. WORK UNIT NUMBER

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)
NATIONAL TAIWAN UNIVERSITY
1, ROOSEVELT RD., SEC. 4
TAIPEI CITY, 10617 TW

8. PERFORMING ORGANIZATION REPORT NUMBER

9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)
AOARD
UNIT 45002
APO AP 96338-5002

10. SPONSOR/MONITOR'S ACRONYM(S)
AFRL/AFOSR IOA

11. SPONSOR/MONITOR'S REPORT NUMBER(S)
AFRL-AFOSR-JP-TR-2018-0027

12. DISTRIBUTION/AVAILABILITY STATEMENT
A DISTRIBUTION UNLIMITED: PB Public Release

13. SUPPLEMENTARY NOTES

14. ABSTRACT
The fundamental goal of this research project is to comprehend complex collective behavior in human society, to set up the foundation of future possible inference and even control of social collective behavior. Two primary technological directions have been fruitfully developed to accomplish such goal: Innovative data analytics to effectively extract critical features from the data, to rapidly recognize the abrupt changes of data, and to identify the network relationship of the data. In addition to synthesized data, the main results have been verified from real Internet data. A new social network model associated with rate rather than traditional graphical properties has been developed to make inference possible, which has been verified by different Internet datasets. In-depth comprehension of information cascade, a collective behavior that agents ignore own signals/observation but follow others, which serves as a foundation of modern social warfare over Internet. We go beyond traditional analytical models of information cascade and develop further views to look into the insights of such
dynamic system to enable methods to tilt collective behavior over the Internet or over human society. A subsequently new security mechanism for low-complexity sensor networks enables trustworthy operation even in the presence of a good number of compromised devices based on social learning and information cascade.

15. SUBJECT TERMS
Social Dynamics, AOARD

16. SECURITY CLASSIFICATION OF:			17. LIMITATION OF ABSTRACT	18. NUMBER OF PAGES	19a. NAME OF RESPONSIBLE PERSON
a. REPORT	b. ABSTRACT	c. THIS PAGE			SERNA, MARIO
Unclassified	Unclassified	Unclassified	SAR	7	19b. TELEPHONE NUMBER *(Include area code)* 315-227-7002

Standard Form 298 (Rev. 8/98)
Prescribed by ANSI Std. Z39.18

Final Report for AOARD Grant FA2386-14-1-4053
"Social Dynamics: Modeling and Inference"

December 1, 2017

PI information:
Professor Kwang-Cheng (K.C.) Chen
Email: ckc@ntu.edu.tw
National Taiwan University
Graduate Institute of Communication Engineering
1, Sec. 4, Roosevelt Road, Taipei, Taiwan 106
(now)
Email: kwangcheng@usf.edu
University of South Florida
Department of Electrical Engineering
4202 E. Fowler Ave
Tampa, FL 33620
Phone: (813) 974-1023

Period of Performance: May 14, 2014 – May 13, 2017

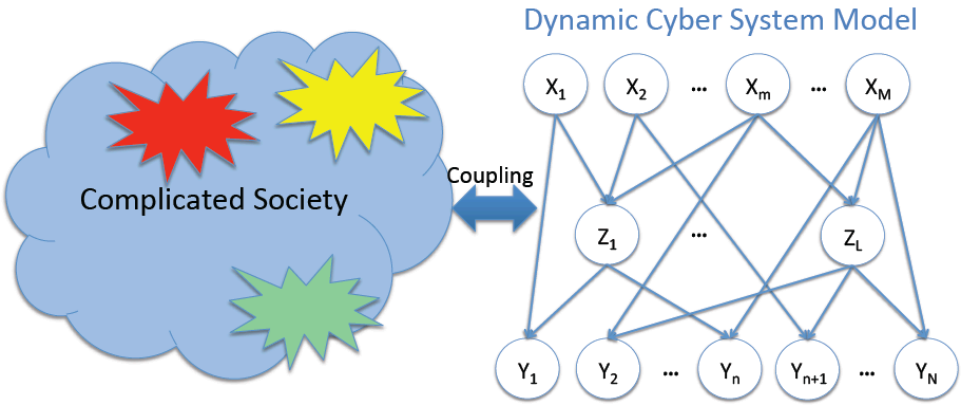
Summary: The fundamental goal of this research project is to comprehend complex collective behavior in human society, to set up the foundation of future possible inference and even control of social collective behavior. Two primary technological directions have been fruitfully developed to accomplish such goal:

- Innovative data analytics to effectively extract critical features from the data, to rapidly recognize the abrupt changes of data, and to identify the network relationship of the data. In addition to synthesized data, the main results have been verified from real Internet data. A new social network model associated with rate rather than traditional graphical properties has been developed to make inference possible, which has been verified by different Internet datasets.
- In-depth comprehension of information cascade, a collective behavior that agents ignore own signals/observation but follow others, which serves as a foundation of modern social warfare over Internet. We go beyond traditional analytical models of information cascade and develop further views to look into the insights of such dynamic system to enable methods to "tilt" collective behavior over the Internet or over human society. A subsequently new security mechanism for low-complexity sensor networks enables trustworthy operation even in the presence of a good number of compromised devices based on social learning and information cascade.

Introduction:

Since the Arabian Spring, linkage between online opinion dynamics and collective behavior in society (particularly political related movements) has been received a lot of interest and considered as a potential tool for social warfare. To complete understand the interplay of online

opinions and collective human behavior, and further to inter (or to predict with statistical confidence) potential human/political/cyber movement(s), are very much wanted goals. The following figure depicts using a large complex dynamic system to represent the complex society. By partial observations of parts of the complicated society (still a big dataset), the model of dynamic networked system can be established to accomplish our ultimate goal.



The goal of this research is to construct new fundamental tools and model(s) toward the ultimate goal. Two primary technologies have been identified to proceed in this investigation:

(a) data analytics to extra networked features (i.e. features of relationship), then develop an effective model enabling inference that was not possible in existing literature
(b) in-depth analytical tools to comprehend information cascade (i.e. rational agents ignoring own observations to follow other agents) and thus collective behavior of agents, such that enables effective influence on agents' behavior.

Above original and fundamental technologies are also useful to develop further resilient and/or proactive defense mechanisms against various cyber-attacks and physical-attacks in cybersecurity and human society.

Experiment: Description of the experiment(s)/theory and equipment or analyses.

Development of innovative theoretical model and methodologies with experimental verifications by realistic Internet data has been executed in this research.

(a) Data analytics toward a new social network model:

- Communication theoretic data analytics: A unique differentiation with common big data analysis for this research lies in the identification of hidden information that implies critical behavior/intention. Such "system" identification can only be derived from open data appearing irrelevant but actually containing tiny amount of information. The methodology based on communication and information theory (thanks to leave at MIT supported by this research) is described in [J1], [C2], and [C4], in which how to select useful data and how to combine useful data for better information extraction are uniquely and systematically developed. A side product is [J5] to introduce machine learning for wireless networks, which has been recognized as IEEE Popular (top 100 papers of most monthly views among 4M papers in the

IEEE Xplore) since publication.

- Similarly, outlier detection (i.e. abrupt changes in behavior or events) and source separation to identify potential root-cause in data analytics have been studied [C3] and [C6].
- New rate-based social network model: Using a real social dataset in Taiwan's PTT, it is confirmed that large social events and online opinion dynamics are closely coupled. Furthermore, such a coupling can be described by network formation and statistical physics (not published). Although the network models for online opinion dynamics have been studied and developed for a long time, existing models are based on network topological information and *impossible* to be applied for effective inference. After all, the rate information from data is used to establish the social network model, which are consistent with existing topological network models. Such a rate-based social network model [C8] is possible for early inference prior to complete network formation, to possibly complete the ultimate goal. [J2], [J3] and [C7] are side products along this social network analysis effort, on sensor and mobile networks.

(b) Information cascade and collective behavior:

- Information cascade: Traditional analytical understanding of information cascade is totally based on Bayesian analysis, which is insufficient to our purpose. We examine information cascade as a dynamic system [C1] and as a social learning mechanism in details [J4]. Furthermore, by incentive seeding and rewiring connections, information cascade can be successfully tilted under various types of random networks (since complete observation of a large-scale social network is impossible for an outsider, he/she can only treat as a random network due to dynamic social network topology) [C5]. A thorough view on information cascade is documented in [J4], an unusually long IEEE journal paper.
- Resilience against attacks: Another benefit of understanding collective behavior is to develop resilient mechanism to against attacks. [C9] uniquely shows that this purpose can be achieved by leverage social learning and information cascade, and illustrated by information fusion in a low-complexity sensor network, even a good number of network nodes being compromised.

Results and Discussion:

There are several amazing possibilities to develop in the near future based on the outcomes of this research.

- Social warfare has been identified by China and Russia. By more verifications on large datasets from social network platforms, our research results are ready to be polished and applied for real implementation of (a) early warning/indicator social/cyber collective behavior, and consequently tilt collective behavior for our purpose (b) early warning of cyber-attacks (i.e. cyber radar) as there must be some unusual behavior in advance, which can be possibly obtained by social network analysis on rate-based network model and communication theoretic data analytics.
- Internet of battle things (IoBT) emerges as an important technology. Instead of completely secure networking and computing, as almost impossible to avoid attacks (either cyber or physical) on low-complexity devices, resilient system design to accommodate the existence of compromised nodes is the way to avoid tremendously

sluggish secure operations to prohibit system performance of original design. Again, compromised behavior has unique behavior patterns that can be identified through social network analysis.

After mid-May 2017 (project ending), research efforts on this project have been continuing to develop a method as an early indicator to rapidly infer collective behavior based on the new social network model, and an analytical investigation on transient formation of information cascade.

List of Publications and Significant Collaborations that resulted from your AOARD supported project:

a) papers published in peer-reviewed journals,

[J1] K.C. Chen, S.L. Huang, L. Zheng, H.V. Poor. (2015, April) Communication Theoretic Data Analytics. *IEEE Journal on Selected Areas in Communications*, vol. 33, no. 4, pp. 663-675.

[J2] T.Y. Chuang, K.C. Chen, H.V. Poor. (2015 May) Information Centric Sensor Network Management via Community Structure. *IEEE Communications Letters*. Vol. 19, no. 5, pp. 767-770.

[J3] H. Hsu, K.C. Chen. (2016, January) A Resource Allocation Perspective on Caching to Achieve Low Latency. *IEEE Communications Letters*. Vol. 20, No. 1, pp. 145-148.

[J4] F. Rosas, J.-H. Hsiao, K.-C. Chen. (2017) A Technical Perspective on Information Cascade. vol. 5, pp. 22605-22633, *IEEE Access*.

[J5] C. Jiang, H. Zhang, Y. Ren, Z. Han, K.C. Chen, L. Hanzo, "Machine Learning Paradigms for Next-Generation Wireless Networks", *IEEE Wireless Communications*, vol. 24, no. 2, pp. 98-105, April 2017.

b) papers published in conference proceedings,

[C1] S.L. Huang, K.C. Chen (2015). Information Cascades in Social Networks via Dynamic System Analysis. *IEEE International Conference on Communications*.

[C2] T.Y. Chuang, C.P. Lu, K.C. Chen (2015). Communication Theoretic Prediction on Networked Data. *IEEE International Conference on Communications*.

[C3] M.C. Wu, K.C. Chen (2015). Outlier Detection in Large-Scale Sensor Network Data Using Shrinkage Estimators. *IEEE GLOBECOM*.

[C4] K.C. Chen, B. Mankir, S.L. Huang, L. Zheng, H.V. Poor (2016). Communication Theoretic Inference on Heterogeneous Data. *IEEE International Conference on Communications*.

[C5] J.H. Hsiao, K.C. Chen (2016). Steering Information Cascades in a Social System by Selective Rewiring and Incentive Seeding. *IEEE International Conference on Communications*.

[C6] M.C. Wu, K.C. Chen (2016). Sparse PCA via Hard Thresholding for Blind Source Separation. *IEEE International Conference on Acoustics, Speech, and Signal Processing*.

[C7] H. Hsu, K.C. Chen (2016). Optimal Cache Time for Epidemic Content Dissemination in Mobile Social Networks. *IEEE International Conference on Communications*.

[C8] T.-H. Fan, K.-C. Chen. (2017) A New Social Network Model of Online Forums. *IEEE GLOBECOM*.

[C9] F. Rosas, K.-C. Chen. (2017) Social Learning Against Data Falsification in Sensor Networks. *International Conference on Complex Networks and Their Applications*. (Award Winning)

Please note that *IEEE GLOBECOM* and *IEEE International Conference on Communications* are the society conferences of the IEEE Communications Society. *IEEE International Conference on Acoustics, Speech, and Signal Processing* is the society conference of the IEEE Signal Processing Society.

c) conference presentations,
as b)

d) manuscripts submitted but not yet published, and
N/A

e) provide a list any interactions with industry or with Air Force Research Laboratory scientists or significant collaborations that resulted from this work.

Light discussion with Dr. Kevin A. Kwiat, AFRL, in a workshop. Our social learning approach on information cascade is highly related to game theoretical study in AFRL. Under the support of Florida Center for Cybersecurity (2017-2018), we have been investigating resilient dynamic multi-agent systems against cyber- and physical-attacks via game theory, particularly part of the system has been compromised.

DD882: As a separate document, please complete and sign the inventions disclosure form. Put n/a in boxes 5 a/b if no inventions resulted from the research.

SF425: Include the Federal Financial Report that has been signed by an official from your business office who can certify that all funds have been expended.

Important Note: If the work has been adequately described in refereed publications, submit a summary that describes all research conducted as described above, and also include summary paragraphs for each cited publication. If possible, submit any reprint(s). If a full report needs to be written, then submission of a final report that is very similar to a full length journal article will be sufficient in most cases.

This document may be as long or as short as needed to give a fair account of the work performed during the period of performance. There will be variations depending on the scope of the work. As such, there is no length or formatting constraints for the final report. Keep in mind the amount of funding you received relative to the amount of effort you put into the report. For example, do not submit a $300k report for $50k worth of funding; likewise, do not submit a $50k report for $300k worth of funding. Include as many charts and figures as required to explain the work.